ISCA Archive
http://www.isca-speech.org/archive

European Conference on Speech Technology
Edinburgh, Scotland, UK
September 1987

THE USE OF COHERENCE COEFFICIENT TO PARAMETERS STATISTICAL DEPENDENCE EVALUATION IN ACOUSTIC ANYLYSIS

J. Psutka*, E. Chan*

ABSTRACT

Frequently at acoustic analysis in the systems of speech recognition the designer has to solve the problem of suitable feature parameters choice. Owing to consequent computer processing, the measurement of a small number of parameters with large information content is required. The shortcoming of the system in this sense is a mutual statistical dependence of measured parameters that often occurs. The paper describes a method of mutual statistical dependence evaluation by means of coherence coefficient. At the coefficient estimation the Parzen weights have been used. Some results of the method employed are presented.

INTRODUCTION

An acoustic analyzer provides at output a series of measured parameters. If a general speech signal with random occurence of consonants and vowels can be at input of the analyzer, the output series may be treated as stationary stochastic process. If number of measured features is p then we have p-dimensional stochastic process

$$X_t = (X_t^1, X_t^2, \ldots, X_t^p) \qquad (1)$$

where t denotes discrete time instant. If the matrix f(q) of spectral densities exists then the coherence coefficient may be established. It plays analogous role as the correlation coefficient. The coherence coefficient conveys magnitude of mutual statistical dependence between the components pairs of the process (1) by means of functions

$$C_{jk}(q) = \frac{|f_{jk}(q)|}{[f_{jj}(q) f_{kk}(q)]^{1/2}}, \quad 0 \leq q \leq \pi . \qquad (2)$$

If denominator in expression (2) equals zero for some values of q then we define $C_{jk}(q)=1$ at these points, hence for the coherence coefficient holds $0 \leq C_{jk}(q) \leq 1$. Thus the mean value of the coherence coefficient near one is due to strong statistical dependence of process components and its value equals zero at statistically independent components.

*University of Technology in Pilsen, Nejedleho sady 14, 306 14 Pilsen, Czechoslovakia

10.21437/ECST.1987-149

COHERENCE COEFFICIENT EVALUATION

It is necessary to determine the diagonal spectral densities $f_{jj}(q)$ and the mutual spectral densities $f_{jk}(q)$ for the coherence coefficient evaluation from expression (2). It can be proved (ref 1) that an impartial estimation of the spectral density provides function

$$I(q) = \frac{1}{2\pi}\sum_{t=1-N}^{N-1} c_t \exp(-itq) \tag{3}$$

where

$$c_t = \frac{1}{N}\sum_{s=1}^{N-t} X_s X_{s+t} \ . \tag{4}$$

Here X_s is random quantity and $c_t = c_{-t}$. In practical cases the estimation according to (3) is not satisfactory since we usually have only one realization of the stochastic process and the dispersion of the quantity (3) does not converge to zero. Therefore somewhat intricate procedure must be employed. For the diagonal spectral density estimation we can take

$$\hat{f}(q) = \int_{-\pi}^{\pi} v(x-q) I(x) dx \ . \tag{5}$$

At this point we assume validity of the expansion into Fourier series

$$v(x) = \sum_{t=-\infty}^{\infty} w_t \exp(itx) \ , \tag{6}$$

so that from (3),(5),(6) we obtain

$$\hat{f}(q) = c_0 w_0 + 2\sum_{t=1}^{N-1} c_t w_t \cos(qt) \ . \tag{7}$$

For the coefficients w_t we have estimates by Parzen (ref 2)

$$w_t = \begin{cases} \frac{1}{2\pi}\left[1 - \frac{6t^2}{m}\left(1 - \frac{t}{m}\right)\right] & \text{for } t = 0, 1, \ldots, \frac{m}{2} \\ \frac{1}{\pi}\left(1 - \frac{t}{m}\right)^3 & \text{for } t = \frac{m}{2}+1, \ldots, m \\ 0 & \text{for } t > m \end{cases} \tag{8}$$

where m is even, $m<N$; recommended is $N/6<m<N/5$. The estimations are computed at points

$$q = \pi n/m, \quad n = 0, 1, \ldots, m \tag{9}$$

By the analogous way we obtain the mutual spectral densities estimations $\hat{f}_{jk}(q)$. For components of the stochastic process (1) we have the estimation

$$I_{jk}(q) = \frac{1}{2\pi N} \sum_{t=1}^{N} X_t^j \exp(-itq) \sum_{s=1}^{N} X_s^k \exp(-isq), \quad -\pi \leq q \leq \pi . \tag{10}$$

If we define the function

$$B_{jk}(t) = \frac{1}{N} \sum_{s=1}^{N-t} X_{s+t} X_s , \quad t = 0, 1, \ldots, N-1 \tag{11}$$

$$B_{jk}(t) = B_{jk}(-t)$$

then the estimation found can be written in the form of

$$\hat{f}_{jk}(q) = \sum_{t=1-N}^{N-1} w_t B_{jk}(t) \exp(-itq). \tag{12}$$

This is generally complex function at j≠k, so that we assume

$$\hat{f}_{jk}(q) = \hat{a}_{jk}(q) + i\hat{b}_{jk}(q). \tag{13}$$

From expressions (12) and (13) comes out

$$\hat{a}_{jk}(q) = w_0 B_{jk}(0) + \sum_{t=1}^{N-1} w_t [B_{jk}(t) + B_{kj}(t)] \cos(qt) , \tag{14}$$

$$\hat{b}_{jk}(q) = \sum_{t=1}^{N-1} w_t [B_{kj}(t) - B_{jk}(t)] \sin(qt) . \tag{15}$$

Again, for the coefficients w_t estimations we have the Parzen weights (8) and the estimations $\hat{f}_{jk}(q)$ are computed at mesh (9). The coherence coefficient estimations are calculated by expression

$$\hat{C}_{jk}(q) = \frac{|\hat{f}_{jk}(q)|}{[\hat{f}_{jj}(q)\hat{f}_{kk}(q)]^{1/2}} . \tag{16}$$

RESULTS

The method usefulness has been confirmed by mutual statistical dependence tests with four components of measured features in the acoustic analysis block of an isolated words recognition system. The parameters under test were short time mean values of the speech signal zero crossings (F1), zero crossings after derivation (F2), short time mean values of signal intensity in frequency bands under 650 Hz

(E1) and above 650 Hz (E2). The features were periodically measured every ten miliseconds. The speech signal was phonetically segmented too, so that coherence coefficient average values for various groups of czech vowels and consonants could be determined, cf Table 1. It can be seen rather higher dependence between the features F1, F2 for constrictives (noise consonants) in comparison with nasals and vowels. This conforms to theory as features F1, F2 can roughly pick up the information about 1st resp. 2nd formant presence in the speech signal. In fact, the method described has been used at finding for separating frequency between the features E1, E2 measuring bands while the measured parameters mutual statistical dependence should be as low as it could.

Table 1 Coherence coefficient average values for speech signal features F1, F2, E1, E2 vs Czech vowels and consonants

	F1-F2	F1-E1	F2-E1	E1-E2
Vowels	0.60	0.38	0.32	0.52
Constrictives	0.90	0.34	0.27	0.33
Explosives	0.67	0.36	0.32	0.50
Nasals	0.57	0.40	0.37	0.44

REFERENCES

1. J.Andel, Statisticka analyza casovych rad (SNTL,Praha,1978)
2. E.Parzen, AMS 28, 329-348 (1957)
3. E.J.Hannan, Time Series Analysis (Wiley,NY,1960)